US Army Corps of Engineers
Construction Engineering
Research Laboratories

AD-A283 923
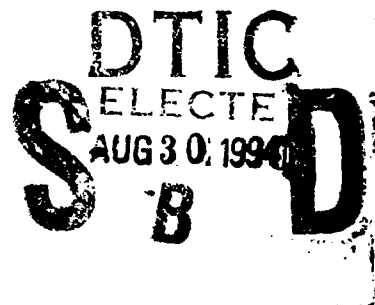
r.mapcalc: An Algebra for GIS and Image Processing

by
Michael Shapiro and James Westervelt

As geographic information systems are enhanced with more robust query capabilities, they evolve from basic spatial data storage, retrieval, and display systems to more powerful spatial modeling systems. Languages that empower a user to design and create spatial models will advance this development even further.

This study describes a map algebra language that can form the foundation for overlay and neighborhood analyses in both raster geographic information systems and image-processing systems. The language combines maps and images in expressions with arithmetic and logical operators and mathematical functions to produce new maps and images. This algebra has been implemented in the Geographical Resources analysis Support System r.mapcalc module. The language syntax is first described and then examples relevant to geographic information systems and image-processing applications are presented.

DTIC
ELECTE
AUG 3 0 1994
S B D

94-27946

94 8 29 230

DESTROY THIS REPORT WHEN IT IS NO LONGER NEEDED

DO NOT RETURN IT TO THE ORIGINATOR

USER EVALUATION OF REPORT

REFERENCE: USACERL Technical Report EC-94/13, *r.mapcalc: An Algebra for GIS and Image Processing*

Please take a few minutes to answer the questions below, tear out this sheet, and return it to USACERL. As user of this report, your customer comments will provide USACERL with information essential for improving future reports.

1. Does this report satisfy a need? (Comment on purpose, related project, or other area of interest for which report will be used.)

_____

_____

_____

2. How, specifically, is the report being used? (Information source, design data or procedure, management procedure, source of ideas, etc.)

_____

_____

3. Has the information in this report led to any quantitative savings as far as manhours/contract dollars saved, operating costs avoided, efficiencies achieved, etc.? If so, please elaborate.

_____

_____

4. What is your evaluation of this report in the following areas?

    a. Presentation: _____

    b. Completeness: _____

    c. Easy to Understand: _____

    d. Easy to Implement: _____

    e. Adequate Reference Material: _____

    f. Relates to Area of Interest: _____

    g. Did the report meet your expectations? _____

    h. Does the report raise unanswered questions? _____

    i. General Comments. (Indicate what you think should be changed to make this report and future reports of this type more responsive to your needs, more usable, improve readability, etc.)

_____

_____

_____

_____

_____   _____

_____

_____

5. If you would like to be contacted by the personnel who prepared this report to raise specific questions or discuss the topic, please fill in the following information.

Name: _____

Telephone Number: _____ _____

Organization Address: _____ __ _ __

6. Please mail the completed form to:

Department of the Army
CONSTRUCTION ENGINEERING RESEARCH LABORATORIES
ATTN: CECER-IMT
P.O. Box 9005
Champaign, IL 61826-9005

REPORT DOCUMENTATION PAGE

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE April 1994	3. REPORT TYPE AND DATES COVERED Final

4. TITLE AND SUBTITLE
r.mapcalc: An Algebra for GIS and Image Processing

5. FUNDING NUMBERS
4A162720
A896
NN-TS2

6. AUTHOR(S)
Michael Shapiro and James Westervelt

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

U.S. Army Construction Engineering Research Laboratories (USACERL)
P.O. Box 9005
Champaign, IL 61826-9005

8. PERFORMING ORGANIZATION REPORT NUMBER

EC-94/13

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Office of the Assistant Chief of Staff for Installation Management (ACS(IM))
ATTN: DAIM-ED-N
600 Army, Pentagon
Washington, DC 20310-0600

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

Copies are available from the National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161.

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

As geographic information systems are enhanced with more robust query capabilities, they evolve from basic spatial data storage, retrieval, and display systems to more powerful spatial modeling systems. Languages that empower a user to design and create spatial models will advance this development even further.

This study describes a map algebra language that can form the foundation for overlay and neighborhood analyses in both raster geographic information systems and image-processing systems. The language combines maps and images in expressions with arithmetic and logical operators and mathematical functions to produce new maps and images. This algebra has been implemented in the Geographical Resources analysis Support System r.mapcalc module. The language syntax is first described and then examples relevant to geographic information systems and image-processing applications are presented.

14. SUBJECT TERMS
geographic information systems   geographic resources analysis support system
spatial modeling system
image processing

15. NUMBER OF PAGES
26

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT SAR

Foreword

Contents

1 INTRODUCTION

Background

Because geographic information systems (GISs) have been enhanced with more robust query capabilities, they have evolved from basic spatial data storage. retrieval, and display systems to more powerful spatial modeling systems. Languages that empower an end-user to design and create spatial models will advance this development even further. Tomlin (1990) described a spatial modeling language for raster data with three major classes of map operations: functions of individual cells (e.g., map recoding and Boolean overlay); functions of cells within neighborhoods (e.g., filters and slope-aspect generation); and functions of cells within regions or zones (e.g., area calculations). This classification could be applied to image-processing operations as well because they are inherently raster-based and employ analysis methods very similar to some GIS operations (Burrough 1986).

A GIS map algebra was needed that would accommodate Tomlin's map classifications, allow image-processing operations and could be implemented in the Geographical Resources Analysis Support System (GRASS). GRASS is a public domain, image-processing and geographic information system originally developed by researchers in the Environmental Division of the U.S. Army Construction Engineering Research Laboratories (USACERL) in Champaign, IL. The GRASS system is used to input, analyze, and output geographic data by users in both military and nonmilitary, public and private agencies, based in North America, Europe, and other parts of the world. It is a component of the Integrated Training Area Management program implemented by USACERL to assist in the management of Army training lands.

Objective

The objective of this research was to develop a map algebra language flexible enough to accommodate operations in Tomlin's first two classes of map operations and that would allow developers and end-users to construct a core of GIS analysis and image-processing operations.

Approach

A map algebra language syntax was first defined and developed and then tested and illustrated with examples relevant to GISs and image-processing applications. The sample geographic data used in the examples came from the Spearfish, SD database–a database distributed with GRASS software.

Mode of Technology Transfer

The map algebra has been implemented in the *r.mapcalc* module of the Geographic Resources Analysis Support System. GRASS is transferred to the field through the following mechanisms: training programs, hands-on experience, a user documentation center, newsletters, institutional structures at the Army and Interagency levels, communication networks, and other forums. GRASS source codes can be obtained via Internet computer software from a file transfer protocol server at ftp moon.cecer.army.mil. Current information is available from the GRASS Information

Center at USACERL, which can be contacted by phone at 217-373-7220, by fax at 217-373-7222, or by e-mail at gic@zorro.cecer.army.mil.

2 LANGUAGE SYNTAX

The *r.mapcalc* algebra employs mathematical operators and functions in expressions involving maps or images that are stored in raster grid-cell format and are two-dimensional matrices of integer values.

The syntax for the algebra is *result=expression*, where *expression* is built using *maps* and *images*, mathematical *operators*, *functions*, and temporary *variables*. The *result* map is produced by evaluating the expression for each cell in the matrix.

For example, the expression:

$$sum = map1 + map2$$

would produce a map *sum* where each cell is the sum of the values of the corresponding cells in *map1* and *map2*.

Maps and Images

Maps and images are database files stored in raster format, i.e., two-dimensional matrices of integer values. The values stored in maps may represent categorical (e.g., soil types) or continuous (e.g., elevation) data, whereas the values stored in images usually represent continuous data (e.g., satellite sensor data). Although the information represented by the values within a map or an image varies, the data format is the same. Therefore, this distinction between map and image will be dropped, and the term *map* will be used for either.

Map naming rules are flexible; however, a map name must not contain any special characters (operators, parentheses, etc.), and it must be distinguishable from numbers. Common names include *elevation, soils, vegetation, roads, band.1*, etc., although the format *12oct89* is also acceptable.

Map names may be followed by a *neighborhood* modifier which specifies a relative offset from the current cell being evaluated. The format is *map[r,c]*, where *r* is the row offset and *c* is the column offset. For example, *map[1,2]* refers to the cell one row below and two columns to the right of the current cell, *map[-2,-1]* refers to the cell two rows above and one column to the left of the current cell, and *map[0,1]* refers to the cell one column to the right of the current cell. This syntax permits the development of neighborhood-type filters within a single map or across multiple maps.

At present, GRASS map files may contain only integer values. However, GRASS does permit floating point values to be associated with each integer value in a map.[1] Maps may be prefixed by the @ modifier which translates the integer values in a map to their associated floating point values.

---

[1]These values are stored in a separate attribute file known as the "category label file."

5

Operators

The operators for the map algebra are:

operator	meaning	type	precedence
*	multiplication	arithmetic	4
/	division	arithmetic	4
%	modulus	arithmetic	4
+	addition	arithmetic	3
−	subtraction	arithmetic	3
==	equal	comparison	2
!=	not equal	comparison	2
<	less than	comparison	2
<=	less than or equal	comparison	2
>	greater than	comparison	2
>=	greater than or equal	comparison	2
&&	and	logical	1
||	or	logical	1

They are applied from left to right, with those of higher precedence applied before those of lower precedence. Parentheses may be used to control the order of evaluation. The arithmetic operations have their usual meanings, except that division by zero equals zero. If both operands are integer, the result is integer, otherwise the result is floating point. The comparisons evaluate to 1 if the comparison is true, otherwise they evaluate to 0. The || operation evaluates to 1 if either operand is nonzero (true), and to 0 otherwise. The && operation evaluates to 1 only if both operands are nonzero (true), and to 0 otherwise.

Functions

Functions perform specific operations on a list of expressions and result in a single operand. The *r.mapcalc* algebra includes the following functions:

abs(x)	$\lvert x \rvert$
exp(x)	e^x
exp(x,y)	x^y
log(x)	$\ln x$
log(x,y)	$\log_y x$
sqrt(x)	$\sqrt{x}$
atan(x)	$\tan^{-1} x$ $[-90°, 90°]$
atan(x,y)	$\tan^{-1} y/x$ $[0°, 360°]$
cos(x)	$\cos x°$
sin(x)	$\sin x°$
tan(x)	$\tan x°$
min(x,y,...)	minimum of $\{x, y, ...\}$
max(x,y,...)	maximum of $\{x, y, ...\}$
float(x)	convert x to floating point
round(x)	round x to nearest integer
int(x)	convert x to largest integer less than or equal to x
if(x)	1, if $x \, != 0$; 0 otherwise
if(x,y)	y, if $x \, != 0$; 0 otherwise
if(x,y,z)	y, if $x \, != 0$; z otherwise
if(x,p,z,n)	p, if $x > 0$; z, if $x == 0$; n otherwise
eval(a,b,...,x)	x (but all arguments are evaluated)

Temporary Variables

Expressions can become complicated and values computed in one part of an expression may need to be computed again in a later part. Such a v   ie can be assigned a temporary *variable* to be used in place of that value in the later part of the expression. For example, *result=(map+2)*(map+2)* may be expressed, using temporary variable x, as *result=(x=map+2)*x*. Additional examples are given on pages 10 and 13.

3   GIS Examples

The *r.mapcalc* algebra supports a variety of GIS operations, some of which are described in the following sections; these, however, do not exhaust the flexibility and power of the language. In the examples, the following maps are used:

| *elevation* | 30-meter digital elevation, range 1066-1840 meters |

| *rushmore* | Camp Rushmore (a fictitious military installation) |

0	outside the installation
1	inside the installation

| *slope* | slope in degrees |

| *vegcover* | vegetation cover |

1	irrigated agriculture
2	rangeland
3	coniferous forest
4	deciduous forest
5	mixed forest
6	disturbed

Map Recoding

One standard GIS operation is the recoding of values in a map. For example, a map of forest cover versus nonforest cover can be created from the *vegcover* map by recoding forest categories to 1, and nonforest categories to 2 as follows:

$$simple.cover \;=\; if(vegcover == 3 \;||\; vegcove \;\; == 4 \;||\; vegcover == 5, \; 1) + \backslash$$
$$if(vegcover == 1 \;||\; vegcover == 2 \;||\; vegcover == 6, \; 2)$$

One adds the results of both *if* functions because each *if* selects a distinct subset of the input data. Note the use of $\backslash$ to indicate that the expression continues on multiple lines.

To make a map of elevations between 1200 and 1375 meters (with elevations outside this range recoded to 0), use:

$$upland = if(elevation >= 1200 \;\&\&\; elevation <= 1375, elevation)$$

Selections

An operation related to map recoding is the selection or identification of cells meeting some specified criteria in one or more maps. Selection differs from recoding in two respects: (1) it may involve more than one map; and (2) the resulting map often has only the values 0 and 1, 1 indicating cells that meet the criteria and 0 indicating those that do not. The logical operators ($\&\&$ and $||$) together with the comparison operators ($==$, $!=$, $<$, $<=$, $>$, and $>=$) provide this selection capability.

For example, to create the map *upland*, with elevations between 1200 and 1375 meters coded as 1, and elevations outside this range coded as 0, use:

$$upland = elevation >= 1200 \;\&\&\; elevation <= 1375$$

8

To create the map *forest*, indicating where forests are found, use:

$$forest = vegcover == 3 \ || \ vegcover == 4 \ || \ vegcover == 5$$

Then, to create the combined map *upland.forest*, indicating where upland forests are found use:

$$upland.forest = forest \ \&\& \ upland$$

Or, more directly, make the map *upland.forest* without first creating *upland* and *forest* by using:

$$upland.forest = (elevation >= 1200 \ \&\& \ elevation <= 1375) \ \&\& \ \backslash$$
$$(vegcover == 3 \ || \ vegcover == 4 \ || \ vegcover == 5)$$

Region Growing

A region-growing operation adds a one-cell border around an area of a map and can be implemented with a simple algorithm. Each zero value in a map is replaced by a nonzero value from the cell to the left, to the right, above, or below.

For example, to add a one-cell border to the Camp Rushmore map, *rushmore*, one can create *rushmore.grow*:

$$rushmore.grow = if(rushmore, rushmore, \backslash$$
$$if(rushmore[0,-1], rushmore[0,-1], \backslash$$
$$if(rushmore[0,1], rushmore[0,1], \backslash$$
$$if(rushmore[-1,0], rushmore[-1,0], \backslash$$
$$if(rushmore[1,0], rushmore[1,0]))))) $$

The border itself, *rushmore.border*, can now be extracted by subtracting the original map from the new map:

$$rushmore.border = rushmore.grow - rushmore$$

Note that a count of just the border cells can be used to approximate the perimeter of the installation.[2]

Slope and Aspect

The neighborhood syntax of the *r.mapcalc* algebra enables a user to calculate a slope gradient (the maximum rate of change in altitude) and its aspect (the compass direction of the slope) from elevation values. The basic formulas (Dozier and Strahler 1983) are:

$$tan(slope) = \sqrt{(\delta z/\delta x)^2 + (\delta z/\delta y)^2}$$
$$tan(aspect) = \frac{\delta z/\delta x}{\delta z/\delta y}$$

---

[2]Cell counting, a regional operation, is not supported by *r.mapcalc*. Cell counts must be determined using another tool (e.g., the GRASS *r.stats* command).

where $\delta z/\delta x$ and $\delta z/\delta y$ are the partial derivatives in the east-west and north-south directions, respectively. Numerical methods can estimate these derivatives (Skidmore 1989). A method given by Horn (1981) is:

$$[\delta z/\delta x]_{y,x} = (z_{y-1,x-1} + 2z_{y,x-1} + z_{y+1,x-1} - z_{y-1,x+1} - 2z_{y,x+1} - z_{y+1,x+1})/8\Delta x$$

$$[\delta z/\delta y]_{y,x} = (z_{y-1,x-1} + 2z_{y-1,x} + z_{y-1,x+1} - z_{y+1,x-1} - 2z_{y+1,x} - z_{y+1,x+1})/8\Delta y$$

where $z_{y,x}$ is the elevation value at row y column x, Δx is the east-west (i.e., column) grid spacing, and Δy is the north-south (i.e., row) grid spacing. This can be expressed more clearly with a matrix of coefficients:

$$\delta z/\delta x = \frac{\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}}{8\Delta X} \qquad \delta z/\delta y = \frac{\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}}{8\Delta Y}$$

The *elevation* map, with its 30-by-30-meter horizontal grid spacing, can be processed with *r.mapcalc* as follows:

```
slope  =  eval(  x =  (elevation[-1,-1] + 2 * elevation[0,-1] + elevation[1,-1]  \
                      -elevation[-1,1] - 2 * elevation[0,1] - elevation[1,1]      \
                      )/(8.0 * 30.0) ,                                            \
                 y =  (elevation[-1,-1] + 2 * elevation[-1,0] + elevation[-1,1]   \
                      -elevation[1,-1] - 2 * elevation[1,0] - elevation[1,1]      \
                      )/(8.0 * 30.0) ,                                            \
                 atan(sqrt(x * x + y * y))                                        \
              )

aspect =  eval(  x =  (elevation[-1,-1] + 2 * elevation[0,-1] + elevation[1,-1]   \
                      -elevation[-1,1] - 2 * elevation[0,1] - elevation[1,1]      \
                      )/(8.0 * 30.0) ,                                            \
                 y =  (elevation[-1,-1] + 2 * elevation[-1,0] + elevation[-1,1]   \
                      -elevation[1,-1] - 2 * elevation[1,0] - elevation[1,1]      \
                      )/(8.0 * 30.0) ,                                            \
                 a =  round(atan(x,y)) ,                                          \
                 if(x||y, if(a,a,360))                                            \
              )
```

Note the use of the temporary variables a, x, and y to capture intermediate results for use in a latter part of the expression.

Slope is calculated using the single-argument version of *atan()*, which produces angles in the range of 0 to 90 degrees. For terrain with low relief, however, where degree values may not give enough information, one could multiply the result by 10 to get 10ths of a degree, or omit the *atan()* function to get *tan(slope)* instead.

Aspect is calculated using the two-argument version of atan(), which produces angles in the range of 0 to 360 degrees. However, because aspect is undefined if both x and y are zero (i.e., flat terrain), additional care is required. The code:

$$a = round(atan(x,y))$$
$$if(x||y, if(a,a,360))$$

produces values from 1 to 360 for cells that have nonzero slope, and a zero value for flat terrain.

Frank (1988) used an alternative technique to compute $\delta z/\delta x$ and $\delta z/\delta y$:

$$\delta z/\delta x = \frac{\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 8 & 0 & -1 & -8 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}}{12\Delta X} \qquad \delta z/\delta y = \frac{\begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & -8 & 0 & 0 \end{bmatrix}}{12\Delta Y}$$

This is expressed in *r.mapcalc* as:

```
slope   = eval(  x =  (elevation[0,-2] + 8 * elevation[0,-1]   \
                       -elevation[0,2] - 8 * elevation[0,1]     \
                       )/(12.0 * 30.0) ,                        \
                 y =  (elevation[-2,0] + 8 * elevation[-1,0]    \
                       -elevation[2,0] - 8 * elevation[1,0]     \
                       )/(12.0 * 30.0) ,                        \
                 atan(sqrt(x * x + y * y))                      \
              )


aspect  = eval(  x =  (elevation[0,-2] + 8 * elevation[0,-1]   \
                       -elevation[0,2] - 8 * elevation[0,1]     \
                       )/(12.0 * 30.0) ,                        \
                 y =  (elevation[-2,0] + 8 * elevation[-1,0]    \
                       -elevation[2,0] - 8 * elevation[1,0]     \
                       )/(12.0 * 30.0) ,                        \
                 a =  round(atan(x,y)) ,                        \
                 if(x||y, if(a,a,360))                          \
              )
```

Hydrologic Simulation

r.mapcalc can construct a simple hydrologic model that iteratively "flows" water across a landscape. A constant amount of water is first deposited in each cell in the landscape. Then, at each time-step, a portion of the water in each cell is drained into its eight surrounding neighbors. The basic logic is:

```
for each cell
    for each neighbor
        let height difference = (cell elevation + water height)
                minus (neighbor elevation + water height)
        if the height difference is positive
        then
            if the cell elevation is greater than
                    the neighbor's elevation + water height
            then
```

> drain out a portion of the water
>
> *otherwise*
>
> > drain out a portion of the height difference
>
> *otherwise*
>
> > *if the neighbor's elevation is greater than*
> >
> > *the cell's elevation + water height*
> >
> > *then*
> >
> > > drain in a portion of the neighbor's water
> >
> > *otherwise*
> >
> > > drain in a portion of the height difference
>
> *end*
>
> *end*

The outer loop is for each cell in the landscape and is done automatically by *r.mapcalc*. The inner loop applies to the eight neighboring cells and must be explicitly coded by the user.

Before running this model, the elevation map must be converted to the same units as the water map and must be filtered to smooth the elevation. In this example, let us assume that the water will be in units of 0.1 inches. To convert elevation from meters to 0.1-inch units:

$$elev = elevation * 393.7$$

The map *elev* can then be smoothed with the following filter:

$$\frac{\begin{bmatrix} 1 & 2 & 1 \\ 2 & 8 & 2 \\ 1 & 2 & 1 \end{bmatrix}}{20}$$

as follows:

$$
\begin{aligned}
elev \;=\; (\quad & elev[-1,-1] + 2 * elev[-1,0] + elev[-1,1] \quad \backslash \\
+ \;\; & 2 * elev[0,-1] + 8 * elev + 2 * elev[0,1] \quad \backslash \\
+ \;\; & elev[1,-1] + 2 * elev[1,0] + elev[1,1] \quad \backslash \\
) \;\; & /20
\end{aligned}
$$

(This filter produces a map with smoother contours. It also tends to create dams at terrain choke points, which will form little lakes during the simulation.)

The model itself is coded as follows:

$$
\begin{aligned}
water = \quad &water + eval(x = elev + water, \\
&if \ (x > (y = elev[-1,0] + water[-1,0])), \\
&\quad -.15 * if \ (elev > y, \quad water, \quad x - y), \\
&\quad .15 * if \ (elev[-1,0] > x, \quad water[-1,0], \quad y - x)) + \\
&if \ (x > (y = elev[1,0] + water[1,0])), \\
&\quad -.15 * if \ (elev > y, \quad water, \quad x - y), \\
&\quad .15 * if \ (elev[1,0] > x, \quad water[1,0], \quad y - x)) + \\
&if \ (x > (y = elev[0,-1] + water[0,-1])), \\
&\quad -.15 * if \ (elev > y, \quad water, \quad x - y), \\
&\quad .15 * if \ (elev[0,-1] > x, \quad water[0,-1], \quad y - x)) + \\
&if \ (x > (y = elev[0,1] + water[0,1])), \\
&\quad -.15 * if \ (elev > y, \quad water, \quad x - y), \\
&\quad .15 * if \ (elev[0,1] > x, \quad water[0,1], \quad y - x)) + \\
&if \ (x > (y = elev[-1,1] + water[-1,1])), \\
&\quad -.10 * if \ (elev > y, \quad water, \quad x - y), \\
&\quad .10 * if \ (elev[-1,1] > x, \quad water[-1,1], \quad y - x)) + \\
&if \ (x > (y = elev[1,1] + water[1,1])), \\
&\quad -.10 * if \ (elev > y, \quad water, \quad x - y), \\
&\quad .10 * if \ (elev[1,1] > x, \quad water[1,1], \quad y - x)) + \\
&if \ (x > (y = elev[1,-1] + water[1,-1])), \\
&\quad -.10 * if \ (elev > y, \quad water, \quad x - y), \\
&\quad .10 * if \ (elev[1,-1] > x, \quad water[1,-1], \quad y - x)) + \\
&if \ (x > (y = elev[-1,-1] + water[-1,-1])), \\
&\quad -.10 * if \ (elev > y, \quad water, \quad x - y), \\
&\quad .10 * if \ (elev[-1,-1] > x, \quad water[-1,-1], \quad y - x)))
\end{aligned}
$$

The resulting map *water* represents the water depth after a single iteration (time-step). Note that the eight sections in this model (beginning with $if(x > ...)$) are similar, each handling a neighboring cell. The first four drain 15 percent of the water to or from the cells above, below, and to either side; the last four drain 10 percent of the water to or from the cells at the diagonal positions.

This model has been applied to a section of the Spearfish database with realistic-looking results. (The data set entitled *spearfish* is a sample database included with the GRASS release and covers the Spearfish, South Dakota vicinity.) Water drains away from the uplands forming temporary streams and ponds. For those wishing to repeat the experiment with GRASS, enter the model code in a file called *water.mapcalc*, then create a controlling shell script:

```
#!/bin/sh
r.mapcalc water=120 # start with 12 inches of water in each cell
d.rast water        # display the water map
i=1
while [ $i != 100 ]
do
        n=1
        while [ $n != 10 ]
```

```
      do
              r.mapcalc < water.mapcalc  # run the simulation
              d.rast water               # display the water map
              n=`expr $n + 1`
      done
      g.copy rast=water,water.$i         # save a snapshot
      i=`expr $i + 1`
done
```

This script will run 1000 iterations. The depth starts at 120 (12 inches, or 30.48 cm, of water) across the entire map. After every tenth iteration, the resultant map is copied to a map with a unique name[3] for rapid display later. One can run the shell script and watch the water "flow" off the slopes of South Dakota's Black Hills.

This model could be improved by considering additional factors such as: (1) *evaporation and transpiration*—in each time-step (iteration), some water will evaporate, which could be modeled either as a constant value or as a simple function of groundcover type; (2) *flow impedance*—the amount of water that can leave a cell in a given time step could be modified using groundcover information; and (3) *ground saturation*—the rate at which soils absorb water during rainfall can be estimated based on soil type and length of exposure to water.

4   IMAGERY EXAMPLES

The *r.mapcalc* algebra also supports many image-processing operations. Examples of these follow; again, they are not exhaustive, but are intended to illustrate the power of the algebra.

Spectral Ratios

Ratio transformation of spectral data is one technique used in the analysis of remotely sensed data (Lillesand and Kiefer 1987). Between-band ratios, which eliminate multiplicative effects, are easily created with *r.mapcalc*:

$$ratio = band.1/band.2$$

Ratios of between-band differences, which eliminate additive effects, are also straightforward:

$$ratio = (band.1 - band.2)/(band.3 - band.2)$$

A normalized vegetation index (*nvi*) for AVHRR data is computed as:

$$nvi = (avhrr.2 - avhrr.1)/(avhrr.2 + avhrr.1)$$

These results could be produced with a very simple algebra that uses only arithmetic operators. The transformed vegetative index (*tvi*) for the Landsat Thematic Mapper illustrates the use of functions in *r.mapcalc*:

$$tvi = 100 * sqrt((tm.4 - tm.3)/(tm.4 + tm.3) + 0.5)$$

---

[3] These maps will be named *water.1, water.2, ... water.100.*

14

Note that 0.5 is added to help ensure that *sqrt*() is not applied to a negative value. This is guaranteed by using the *max*() function:

$$tvi = 100 * sqrt(max(0, \ (tm.4 - tm.3)/(tm.4 + tm.3) + 0.5))$$

If the results are to be in the range 0 to 255, this ratio will produce such results:

$$ratio = 255.0/90.0 * atan(float(tm.1)/float(tm.2))$$

The *float*() function ensures that the division is floating point, not integer, division. Because the output map can hold only integer values, *r.mapcalc* will round the results to integers before they are stored in the map.

Spatial Filters

A common image-processing operation is local-neighborhood filtering. A small window is moved over an image, and the center pixel is replaced by a combination of all the pixels in the window.

Low-frequency filters (also called low-pass filters) deemphasize high spatial frequencies and involve computing a weighted average of all pixels in the window. A simple average, which smoothes the image, is based on the filter:

$$\frac{\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}}{9}$$

This is expressed with *r.mapcalc* as:

```
ave3x3   =   (   image[-1,-1] + image[-1,0] + image[-1,1]   \
             +   image[0,-1] + image[0,0] + image[0,1]       \
             +   image[1,-1] + image[1,0] + image[1,1]       \
             )   / 9
```

Use high-frequency filters (high-pass filters) to emphasize high spatial frequencies and for edge enhancement. The technique is the same as for low-frequency filters, except that some of the weights are negative. One such filter is:

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

This is expressed with *r.mapcalc* as:

```
high3x3   =   (   -image[-1,-1] - image[-1,0] - image[-1,1]   \
                  -image[0,-1] + 9 * image[0,0] - image[0,1]   \
                  -image[1,-1] - image[1,0] - image[1,1]       \
              )
```

The Sobel filter is a nonlinear edge-enhancement filter:

$$\sqrt{\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}^2 + \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}^2}$$

It is similar to the method to determine slope values from elevation values and is expressed in *r.mapcalc* as:

```
sobel3x3 = eval(   x =   image[-1, 1] + 2 * image[0, 1] + image[1, 1]          \
                        -image[-1, -1] - 2 * image[0, -1] - image[1, -1] ,    \
                   y =   image[-1, -1] + 2 * image[-1, 0] + image[-1, 1]       \
                        -image[1, -1] - 2 * image[1, 0] - image[1, 1] ,        \
                  sqrt(x * x + y * y)                                          \
              )
```

Another type of filter is an adaptive one, which replaces the center pixel only if some criteria are met. Eliason and McEwen (1990) describe two whereby the center pixel is replaced by the average value in the neighborhood if the difference between them exceeds a threshold. One filter sets the threshold to a multiple of the statistical variance in the neighborhood. The other sets the threshold to a multiple of the statistical variance in the entire image. Both filters can be coded with *r.mapcalc*; the variance within the entire image is a regional calculation and must first be computed by another command[4] and then inserted into the *r.mapcalc* code. The first filter is illustrated.

Let P be the value of the center pixel, s the sum of all values in the neighborhood, n the number of pixels in the neighborhood, and ss the sum of the squares of the values in the neighborhood. The average is given by:

$$\mu = \frac{s}{n}$$

and the variance is given by:

$$\sigma^2 = \frac{ss}{n} - \mu^2$$

The center pixel is replaced if:

$$(P - \mu)^2 > C\sigma^2$$

where C is usually between 1.0 and 4.0. The *r.mapcalc* code for a 3x3 window with C set to 2.25 is:

---

[4]The GRASS *r.stats* command could be used to compute the variance across the entire image.

$$filter = eval(\quad s \quad = \quad image[-1,-1] + image[-1,0] + image[-1,1] + \backslash$$
$$image[0,-1] + image[0,0] + image[0,1] + \qquad \backslash$$
$$image[1,-1] + image[1,0] + image[1,1] \ , \qquad \backslash$$

$$ss \quad = \quad image[-1,-1] * image[-1,-1] + \qquad \backslash$$
$$image[-1,0] * image[-1,0] + \qquad \backslash$$
$$image[-1,1] * image[-1,1] + \qquad \backslash$$
$$image[0,-1] * image[0,-1] + \qquad \backslash$$
$$image[0,0] * image[0,0] + \qquad \backslash$$
$$image[0,1] * image[0,1] + \qquad \backslash$$
$$image[1,-1] * image[1,-1] + \qquad \backslash$$
$$image[1,0] * image[1,0] + \qquad \backslash$$
$$image[1,1] * image[1,1] \ , \qquad \backslash$$

$$ave \quad = \quad s/9.0 \ , \qquad \backslash$$
$$var \quad = \quad ss/9.0 - ave * ave \ , \qquad \backslash$$
$$x \quad = \quad image - ave \ , \qquad \backslash$$
$$if(x * x > 2.25 * var, ave, image) \qquad \backslash$$
$$)$$

Radiometric Calibration

Radiometric calibration converts digital values recorded by a remote sensing system to radiance values at the top of the atmosphere, possibly followed by a conversion to surface reflectance values. Radiance values can be computed using the following formula (Chavez 1989):

$$rad_i(x,y) = (dn_i(x,y) - offset_i)/gain_i$$

where *rad* is the radiance for band i at pixel x,y; dn is the value recorded by the sensors; and *offset* and *gain* are the sensor offset and gain for band i. This formula is easily handled by *r.mapcalc*. For example, using the values for the offset and gain as reported by Chavez for October 3, 1988 Thematic Mapper data, radiance images are constructed as follows:

$$rad.1 \quad = \quad (tm.1 - 2.4899)/16.5993$$
$$rad.2 \quad = \quad (tm.2 - 2.3871)/8.5104$$
$$rad.3 \quad = \quad (tm.3 - 1.4815)/12.4074$$
$$rad.4 \quad = \quad (tm.4 - 1.8418)/12.2790$$
$$rad.5 \quad = \quad (tm.5 - 3.4240)/92.5292$$
$$rad.7 \quad = \quad (tm.7 - 2.6323)/175.4878$$

The radiance formula for the SPOT sensor appears in a slightly modified form (USA SPOT Image Corporation 1989):

$$rad_i(x,y) = dn_i(x,y)/A_i$$

where A_i are the absolute calibration coefficients. The May 27, 1989 SPOT multispectral image in the Spearfish database can be converted to radiance, using *r.mapcalc*, as follows:

$$rad.1 \quad = \quad spot.ms.1/1.05586$$
$$rad.2 \quad = \quad spot.ms.2/1.12140$$
$$rad.3 \quad = \quad spot.ms.3/0.97244$$

Surface reflectance calculations are more involved because the formulas must incorporate atmospheric effects. There are various methods for modeling these effects (Richards 1986, Price 1987, Chavez 1989); a formula given by Chavez is:

$$R_i(x,y) = \frac{\pi dist^2[rad_i(x,y) - haze_i]}{E_i slope(x,y) \cdot sun \cdot mhaze_i}$$

where R is the surface reflectance image, rad is the radiance image, $slope$ is a slope map, $dist$ is the Earth-Sun distance, $haze$ and $mhaze$ are the additive and multiplicative atmospheric haze factors, E is exoatmospheric spectral irradiance, and sun is the sun elevation at the time the image was captured. C   ing the nonmap factors into single constants, the formula can be written in the form:

$$R_i(x,y) = \frac{A_i[rad_i(x,y) - B_i]}{slope(x,y)}$$

which, after substituting the appropriate values for the constants A and B, is expressed with r.mapcalc as:

$$R.1 = A_1 * (rad.1 - B_1)/slope$$
$$R.2 = A_2 * (rad.2 - B_2)/slope$$
$$R.3 = A_3 * (rad.3 - B_3)/slope$$

.

.

.

Principal Components

Principal components analysis involves transforming a set of correlated variables into a new, uncorrelated set. The variables are the band files from a multispectral sensor. The transformation is a linear combination of the band data:

$$component_i = \sum_{j=1}^{N} w_{ij} * band_j \quad i = 1, \ldots, N$$

where N is the number of band files and w_{ij} are the eigenvectors for the between-band covariance matrix. Although neither the covariance computation nor the determination of the eigenvectors can be accomplished using the map algebra, the components can be computed. For example, suppose the three multispectral bands from a SPOT image had the following covariance matrix:

$$\begin{bmatrix} 462.88 & 480.41 & 281.76 \\ 480.41 & 513.02 & 278.92 \\ 281.76 & 278.91 & 336.33 \end{bmatrix}$$

A set of eigenvectors for this matrix (in decreasing order of spectral variance) is:

vector$_1$	21.24	22.15	14.77
vector$_2$	2.91	4.46	−10.87
vector$_3$	1.82	−1.62	−0.18

To compute the second component with r.mapcalc:

$$pc.2 = 2.91 * spot.ms.1 + 4.46 * spot.ms.2 - 10.87 * spot.ms.3$$

18

Merging Panchromatic Data With Multispectral Data

The French SPOT satellite produces a 10-meter panchromatic image plus three spectral bands with 20-meter resolution. Landsat Thematic Mapper produces six spectral bands with 30-meter resolution (as well as a 120-meter resolution thermal band). A sharpened color image can be produced by merging the high-resolution panchromatic image with the lower resolution spectral bands. With the assumption that the spectral data have been registered and resampled to the 10-meter panchromatic data, a number of techniques have been used to perform the merger, most of which can be done with the map algebra.

Welch and Ehlers (1987) described two direct methods:

$$M_i' = a_i(w_1 M_i + w_2 P) + b_i$$
$$M_i' = a_i(M_i P)^{\frac{1}{2}} + b_i$$

where M_i is spectral band i, P is the panchromatic band, w_1 and w_2 are weights, and a_i and b_i are chosen to make the results fall within the range 0 to 255. Both these methods can be implemented using the *r.mapcalc* algebra.

A simple average, based on the first method, is coded as follows:

$$merge.1 = (spot.ms.1 + spot.pan)/2$$
$$merge.2 = (spot.ms.2 + spot.pan)/2$$
$$merge.3 = (spot.ms.3 + spot.pan)/2$$

Or a combination of both methods can be used:

$$merge.1 = sqrt(spot.pan * spot.ms.1)$$
$$merge.2 = sqrt(spot.pan * spot.ms.2)$$
$$merge.3 = (spot.pan + 3 * spot.ms.3)/4$$

A third method transforms three selected bands from red, green, and blue color space to intensity, hue, and saturation color space. The intensity is replaced by the panchromatic image, and the new intensity, hue, and saturation combination is transformed back into red, green, and blue.

Intensity, Hue, Saturation

Conversion from red, green, and blue colors to intensity, hue, and saturation involves formulas that can be represented using the *r.mapcalc* algebra. The following is based on the hue, saturation, and value (HSV) algorithm found in Foley and Van Dam (1984). Assuming that R, G, and B represent the red, green, and blue components of an image, the value (V), saturation (S),

and hue (H) can be formulated as follows:

$$
\begin{aligned}
V \;&=\; max(R,G,B) \\
S \;&=\; 255.0 * (V - min(R,G,B))/V \\
H \;&=\; if(S, eval(\qquad\qquad\qquad\qquad\quad \backslash \\
&\qquad m \;=\; float(min(R,G,B)) , \quad \backslash \\
&\qquad\;\; r \;=\; (V-R)/(V-m) , \qquad \backslash \\
&\qquad\;\; g \;=\; (V-G)/(V-m) , \qquad \backslash \\
&\qquad\;\; b \;=\; (V-B)/(V-m) , \qquad \backslash \\
&\qquad\;\; h \;=\; if(R == V, b-g)+ \qquad \backslash \\
&\qquad\qquad\;\; if(G == V, 2+r-b)+ \;\backslash \\
&\qquad\qquad\;\; if(B == V, 4+g-r) , \;\; \backslash \\
&\quad if(h <= 0, h+6, h) * 60))
\end{aligned}
$$

Although saturation is normally defined in the range from 0 to 1, it is multiplied here by 255 to retain more information after conversion to integer. Hue will be in the range of 0 to 360, with 0 representing undefined hues (i.e., white, black, or gray).

The inverse transformation is:

$$
\begin{aligned}
R \;=\; eval(&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \backslash \\
s \;&=\; S/255.0 , \qquad\qquad\qquad\qquad\qquad\qquad \backslash \\
h \;&=\; if(H >= 360, H - 360, H)/60.0 , \qquad\quad \backslash \\
i \;&=\; int(h) , \qquad\qquad\qquad\qquad\qquad\qquad\quad \backslash \\
f \;&=\; h - i , \qquad\qquad\qquad\qquad\qquad\qquad\quad \backslash \\
p \;&=\; V * (1 - s) , \qquad\qquad\qquad\qquad\qquad\;\; \backslash \\
q \;&=\; V * (1 - s * f) , \qquad\qquad\qquad\qquad\quad \backslash \\
t \;&=\; V * (1 - s * (1 - f)) , \qquad\qquad\qquad\; \backslash \\
&if(i == 0, V) + if(i == 1, q) + if(i == 2, p)+ \;\; \backslash \\
&if(i == 3, p) + if(i == 4, t) + if(i == 5, V))
\end{aligned}
$$

$$
\begin{aligned}
G \;=\; eval(&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \backslash \\
s \;&=\; S/255.0 , \qquad\qquad\qquad\qquad\qquad\qquad \backslash \\
h \;&=\; if(H >= 360, H - 360, H)/60.0 , \qquad\quad \backslash \\
i \;&=\; int(h) , \qquad\qquad\qquad\qquad\qquad\qquad\quad \backslash \\
f \;&=\; h - i , \qquad\qquad\qquad\qquad\qquad\qquad\quad \backslash \\
p \;&=\; V * (1 - s) , \qquad\qquad\qquad\qquad\qquad\;\; \backslash \\
q \;&=\; V * (1 - s * f) , \qquad\qquad\qquad\qquad\quad \backslash \\
t \;&=\; V * (1 - s * (1 - f)) , \qquad\qquad\qquad\; \backslash \\
&if(i == 0, t) + if(i == 1, V) + if(i == 2, V)+ \;\; \backslash \\
&if(i == 3, q) + if(i == 4, p) + if(i == 5, p))
\end{aligned}
$$

$$
\begin{aligned}
B \;=\; eval(\; & \\
s \;=\;& S/255.0\,, & \backslash \\
h \;=\;& if(H >= 360, H - 360, H)/60.0\,, & \backslash \\
i \;=\;& int(h)\,, & \backslash \\
f \;=\;& h - i\,, & \backslash \\
p \;=\;& V*(1-s)\,, & \backslash \\
q \;=\;& V*(1-s*f)\,, & \backslash \\
t \;=\;& V*(1-s*(1-f))\,, & \backslash \\
& if(i == 0, p) + if(i == 1, p) + if(i == 2, t) + & \backslash \\
& if(i == 3, V) + if(i == 4, V) + if(i == 5, q))
\end{aligned}
$$

5   CONCLUSION

Providing a map algebra for manipulation of raster map data results in a flexible geographic information system. Given appropriate raster data layers, the type and number of potential data manipulations are virtually limitless. For many applications, users are freed from software limits and are bound only by their ability to employ the algebra. *r.mapcalc* supports three levels of usage: (1) as a resource for users who need to perform specific algebraic functions that are not provided by other GRASS programs; (2) as a foundational tool for advanced GIS users to develop a limitless set of image and map analysis functions; and (3) as a resource for programmers and developers to use in building new functions.

r.mapcalc is perhaps most important at the intermediate level, as a tool for advanced GIS users. A common implementation model for a GIS in larger organizations is one or two sophisticated users supporting numerous casual users. In this context, *r.mapcalc* becomes a language that opens an array of possibilities for spatial analysis and provides a means to develop macros to support routine operations used by less sophisticated users.

6 REFERENCES

Burrough, P. A., *Principles of Geographical Information Systems for Land Resources Assessment*, Monograph on Soils and Resources Survey No. 12 (Oxford University Press, 1986).

Chavez, Pat S., Jr., "Radiometric Calibration of Landsat Thematic Mapper Multispectral Images," *Photogrammetric Engineering and Remote Sensing*, Vol 55, No. 9 (1989), pp 1285-1294.

Dozier, J., and A. H. Strahler, "Ground Investigations in Support of Remote Sensing," in *Manual of Remote Sensing*, Vol 1, edited by R. N. Colwell (American Society of Photogrammetry, 1983).

Eliason, Eric M., and Alfred S. McEwen, "Adaptive Box Filters for Removal of Random Noise from Digital Images," *Photogrammetric Engineering and Remote Sensing*, Vol 56, No. 4 (1990), pp 453-458.

Foley, James D., and Andries Van Dam, *Fundamentals of Interactive Computer Graphics* (Addison-Wesley, 1984).

Frank, Thomas D., "Mapping Dominant Vegetation Communities in the Colorado Rocky Mountain Front Range with Landsat Thematic Mapper and Digital Terrain Data," *Photogrammetric Engineering and Remote Sensing*, Vol 54, No. 12 (1988), pp 1727-1734.

Horn, B. K. P., "Hill Shading and the Reflectance Map," *Proceedings of the I.E.E.E.*, No. 69 (1981).

Lillesand, Thomas M., and Ralph W. Kiefer, *Remote Sensing and Image Interpretation*, 2d ed. (John Wiley & Sons, 1987).

Price, John C., "Calibration of Satellite Radiometers and the Comparison of Vegetation Indices," *Remote Sensing of Environment*, No. 21 (1987), pp 15-27.

Richards, John A., *Remote Sensing Digital Image Analysis* (Springer-Verlag, 1986).

Skidmore, Andrew K., "A Comparison of Techniques for Calculating Gradient and Aspect from a Gridded Elevation Model," *International Journal of Geographical Information Systems*, Vol 3, No. 4 (1989), pp 323-334.

SPOT User's Handbook, Vol 2 (USA SPOT Image Corporation, 1989), pp 2-49.

Tomlin, C. Dana, *Geographic Information Systems and Cartographic Modeling* (Prentice-Hall, 1990).

Welch, R., and M. Ehlers, "Merging Multiresolution SPOT HRV and Landsat TM Data," *Photogrammetric Engineering and Remote Sensing*, Vol 53, No. 3 (1987), pp 301-305.

USACERL Distribution